# In vivo characterization of FACT complex proteins in *C. elegans*

Honors Research Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with Honors Research Distinction

Aislinn Latham

Department of Animal Sciences

The Ohio State University

2018

Project Advisor: Helen Chamberlin, PhD, Department of Molecular Genetics

The Ohio State University

**ABSTRACT**

FACT (facilitates chromatin transcription) is a protein complex involved in transcriptional regulation, DNA replication, and DNA repair processes in eukaryotic organisms. FACT allows transcriptional enzymes to access the DNA for transcription and translation into proteins and is highly conserved across species. However, there is limited knowledge about FACT complexes in multicellular organisms. The objective of this study is to discover more information about the composition and expression of the FACT complex and its role in multicellular organisms by tagging the genes that code for the FACT subunits in order to view the subsequent proteins in the nematode *Caenorhabditis elegans.* The complex is composed of two proteins: SSRP1 and SPT16. In *C. elegans*, two genes (*hmg-3* and *hmg-4*) encode SSRP1 proteins, and one gene (*spt-16*) encodes SPT16. In order to determine the presence of the target subunits in *C. elegans*, sequences encoding a fluorescent tag were inserted into each gene using CRISPR-mediated genome editing. Using fluorescent microscopy conditions, the target proteins were visible *in vivo*, revealing the areas of the organism where the subunits were present. The location of the product in the cells of *C. elegans* differed between the HMG-3 and HMG-4 proteins. HMG-3 had somatic presence during the embryonic stage, but in later larval stages the product was confined to the germline. In contrast, HMG-4 and SPT-16 were present in the somatic and germ cells throughout the developmental process of the organisms. The FACT complex's functions in DNA regulation and repair make it a potential target for cancer and HIV research. Gaining information about the expression of the FACT complex's proteins at different life stages in a multicellular eukaryote creates a base from which future research can draw. Future research may include characterizing the interactions between protein subunits, and determining the number of each subunit protein in the complex.

## INTRODUCTION

Transcriptional processes are essential for normal function and growth of a cell. In every eukaryotic organism, there exists a protein complex important to the transcription of the nucleosome called the FACT (facilitates chromatin transcription) complex. This complex is made up of two protein subunits named SSRP1 (structure-specific recognition protein 1) and SPT16 (suppressor of Ty16) (Winkler and Luger, 2011). These subunits have roles as histone chaperones during transcription, and are involved with transcription, regulation, and repair of DNA, including tumor suppression (Winkler and Luger, 2011). Although the FACT complex is found in a wide number of species and has multiple roles critical for healthy cell division, there has not been much work done on the regulation of the complex in multicellular organisms. By tagging genes responsible for the formation of the subunits in the model organism *Caenorhabditis elegans*, it will be possible to further analyze the regulation of these genes in multicellular organisms. Studying the regulation of these genes and the composition of the FACT complex through observation of gene interaction in *C. elegans* could lead to important information about normal cellular processes as well as abnormal division processes such as tumor growth.

The FACT complex is found in many organisms, including humans and domesticated animals. It has an important role in many processes involving cell division, including transcription, replication, and DNA repair. In transcription, the FACT complex acts as a histone chaperone, which works to unravel tightly wound chromatin and reorganize individual nucleosomes for RNA polymerases to access (Winkler and Luger, 2011). Additionally, the FACT complex may play a role in nucleosome reassembly following the passage of RNA polymerase and stabilizing nucleosomes (Winkler and Luger, 2011). Chromatin regulators such as the FACT complex are necessary for normal cell division and normal cell cycle processes. For instance, inhibition of the

F55A3.3 gene that codes for the SPT16 subunit delayed normal cell processes (Kruger et al., 2015). However, most of the work done to discover the FACT complex's role in normal cell processes has used single-celled yeast as the model organism. There is a lack of information about the regulation of FACT complexes in multi-cellular organisms.

Normal cell cycle processes, growth, and division are not the only factors affected by the regulation of the FACT complex. Studies of yeast FACT have shown it to have a role in activating a tumor-suppressing gene in response to DNA damage (Winkler and Luger, 2011). The subunits of the FACT complex have also been shown to suppress transcription of HIV-1 and promote viral latency (Huang et al., 2016). Because of these properties, the FACT complex and it constituent subunits could be promising targets for future treatment and identification of HIV infections, cancers, and other abnormal cell growth conditions. In order to take advantage of the FACT complex's many roles in the cell and its potential for cancer treatment and HIV eradication, more information on its specific composition and activity in multicellular organisms needs to be collected.

This project included two aims:

Aim 1: Insert the green fluorescent protein (GFP) gene into genes encoding parts of the FACT complex: *hmg-3*, *hmg-4*, and *spt-16*.

Aim 2: Observe the abundance and localization of the protein products from *hmg-3*, *hmg-4*, and *spt-16* in *C. elegans* through observation of tagged proteins under a fluorescent microscope and Western Blot procedure.

## MATERIALS AND METHODS

### Aim 1: Creation of Plasmids for Genome Editing

CRISPR-mediated genome editing was used to tag the genes in the *C. elegans* genome. This procedure requires two critical components: a guide RNA (gRNA) and a repair template. For each gene (*hmg-3*, *hmg-4*, and *spt-16*) gRNA was designed and amplified via PCR. The gRNA was then ligated into an empty Cas-9-gRNA plasmid, pDD162 (from addgene.org). This plasmid with the inserted gRNA was used to transform competent E. coli cells. The cells were transformed through a series of temperature-shocking steps before being plated on LB and Carb plates and incubated at 37°C. From the colonies on the plates, cultures were taken and incubated. The plasmids were prepared and collected following the protocol for the Monarch Plasmid DNA Miniprep Kit (New England BioLabs, n.d.). The final product was analyzed for successful insertion using a diagnostic digest as well as verification by sequencing. The gRNA in combination with the CRISPR-associated protein 9 (CAS9) was used to create a double strand break at a specific location in the host DNA for insertion of the GFP gene (Dickinson and Goldstein, 2016). In order to insert the gene into the double strand break for expression, a repair template was created and verified by sequencing. Two regions of homology arms that flank the double-strand break were designed and amplified via PCR. Following confirmation of the successful amplification of the homology arms, which was determined using gel electrophoresis, the homology arms were inserted via Gibson assembly into repair vector pDD282 (from addgene.org) which contained GFP, a self-excising cassette, and 3xFLAG tag. Sequences of potential insertion candidates were compared to the template for a successful insertion using NCBI's nucleotide BLAST sequence alignment program (NCBI, n.d.). A Myc tag was inserted into *C. elegans* strains using a similar protocol to the one for FLAG tag insertion for use in

Western Blot procedures. The repair template plasmid vector for the Myc tag was pDD286 (from addgene.org), which contained tagRFP-T, a self-excising cassette, and 3xMyc tag.

**Aim 1: Generation of tagged genes**

The gRNA, Cas9, the repair template, which included GFP, a self-excising cassette, and 3x FLAG tag were injected into *C. elegans.* Two to three days after injection, plates were treated with hygromycin and three to seven days after treatment with hygromycin were analyzed for the presence of hygromycin-resistant worms with an abnormal rolling-type motion. Presence of hygromycin-resistant rollers indicates successful insertion (Dickinson and Goldstein, 2016). After plates with successful insertion were identified, the self-excising cassette was removed via heat-shock treatment, leaving behind the GFP and 3x FLAG tag (Dickinson and Goldstein, 2016). Insertion of the sequences at the correct location was verified using PCR. Following insertion of the tag and removal of the self-excising cassette, the worms were viewed under fluorescent microscopy conditions to determine the location of the protein products. One strain each was confirmed for *hmg-3::gfp* and *spt-16::gfp,* and two strains were confirmed for *hmg-4::gfp*.

**Aim 2: Microscopy**

Fluorescent images were taken on a Zeiss Axioplan 2 microscope. Larvae were immobilized on agar pads of 3.5% noble agar in water with 10nM sodium azide added to the agar. Early stage embryos were dissected from young adults and later stage embryos were collected from mixed-stage plates. Larval images were taken at 40x or 100x, and embryos were imaged at 100x magnification.

### Aim 2: Western Blot

Western blot protocol was adapted from Michael Koelle's Western Blot protocol (n.d.). Proteins were separated using a 12% polyacrylamide resolving gel and then transferred to a nitrocellulose membrane overnight. The membrane was blocked with a 5x blotto solution of dry milk and TBS/Tween, then washed with TBS/Tween before the appropriate antibody was added to TBS/Tween and washed over the membrane. For FLAG-tagged strains an anti-FLAG tag antibody catalog number A9469 from Sigma-Aldrich at a 1:1000 dilution was used. The Western Blot on Myc-tagged strains will utilize anti-Myc tag antibody catalog number ab81654 from Abcam at a 1:5000 dilution. Alkaline phosphatase developer solution was used to develop the bands on the membrane. Estimated weights of the expected proteins were compared to a protein ladder to confirm the presence of each protein.

## RESULTS

### Aim 1: Creation of Plasmids for Genome Editing

There was successful insertion of gRNA and repair templates into the appropriate plasmids. Insertions of gRNA into pDD162 were confirmed via diagnostic digest and gel (Figure 1), along with DNA sequencing (Figure 2). A successful insertion of gRNA into the plasmids contains the Cas-9 promoter region and transcriptional region for Cas-9, as well as the created gRNA (Figure 3). The insertion of the repair template and homology arms into pDD286 was successful and confirmed via gel electrophoresis (Figure 4).

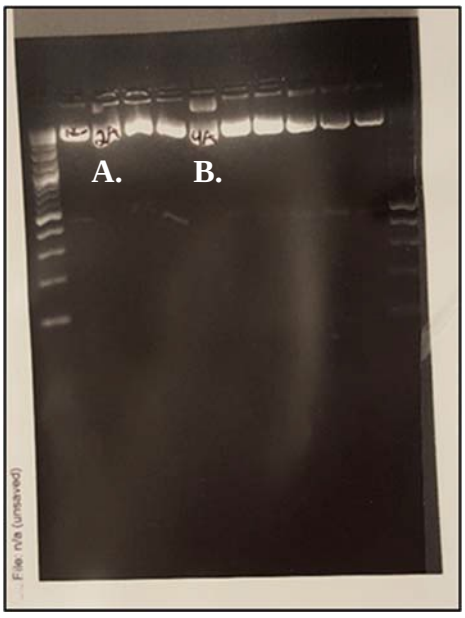


**Figure 1. gRNA diagnostic digest to confirm successful insertion of gRNA into plasmid.** Each sample with a successful insertion of the gRNA was expected to have two fragments. The samples labelled *A.* and *B.* did not have the smaller weight fragment as can be seen in the other samples and therefore did not contain the insertion. Out of the nine samples, seven were confirmed to have the gRNA insertion into the plasmid.

Sequence ID: Query_52503 Length: 300 Number of Matches: 1

Range 1: 107 to 169 Graphics ▼ Next Match ▲ Previous Match

| Score | Expect | Identities | Gaps | Strand |
|---|---|---|---|---|
| 117 bits(63) | 7e-32 | 63/63(100%) | 0/63(0%) | Plus/Minus |

```
Query  1    CCTATTGCGAGATGTCTTGTACTTTCTGCGTTTGTGCGAGTTTTAGAGCTAGAAATAGCA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  169  CCTATTGCGAGATGTCTTGTACTTTCTGCGTTTGTGCGAGTTTTAGAGCTAGAAATAGCA  110

Query  61   AGT  63
            |||
Sbjct  109  AGT  107
```

**Figure 2. Alignment sequence for successful gRNA insertion.** The query represents the expected sequence for a successful gRNA insertion. The subject line represents the sample sequence from the generated plasmids. The gRNA insertion was successful as the query and subject matched with no deletions or base pair differences.

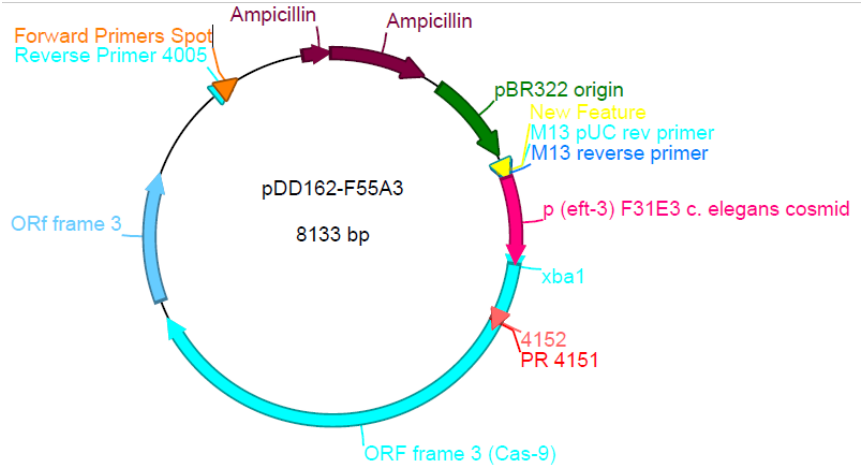


**Figure 3. Gene map for *spt-16*.** Representation of regions within the pDD162 vector with F55A3.3 (also known as *spt*-16) insertion. The plasmid includes a promoter region to initiate transcription of what will become the cas-9 protein and another promoter region for the gRNA transcription.

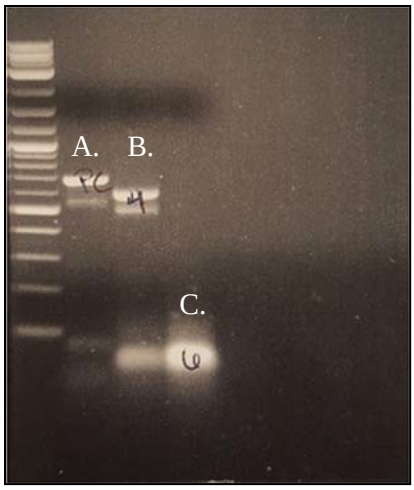


**Figure 4. PCR confirmation for successful assembly of repair template into vector pDD282.** *A.* Positive control for the assembly of the repair template into the vector. *B.* Successful insertion of repair template. *C.* Unsuccessful attempt at insertion of the repair template.

**Aim 1: Generation of Tagged Genes**

Following plasmid creation, transformed cells were confirmed for complete assembly of gRNA and tag (Figure 6). Successfully tagged genes were placed into injection mixes for use in *C. elegans* (Figure 5). The injection mixes for GFP FLAG-tagged genes contains 10 ng/µL of repair template plasmid, 50 ng/µL Cas-9-gRNA plasmid, and 15 ng/µL myo-2 GFP. The injection mix for Myc-tagged *hmg-3* contains 10 ng/µL of repair template plasmid, 50 ng/µL Cas-9-gRNA plasmid, and 25 ng/µL myo-2 GFP.

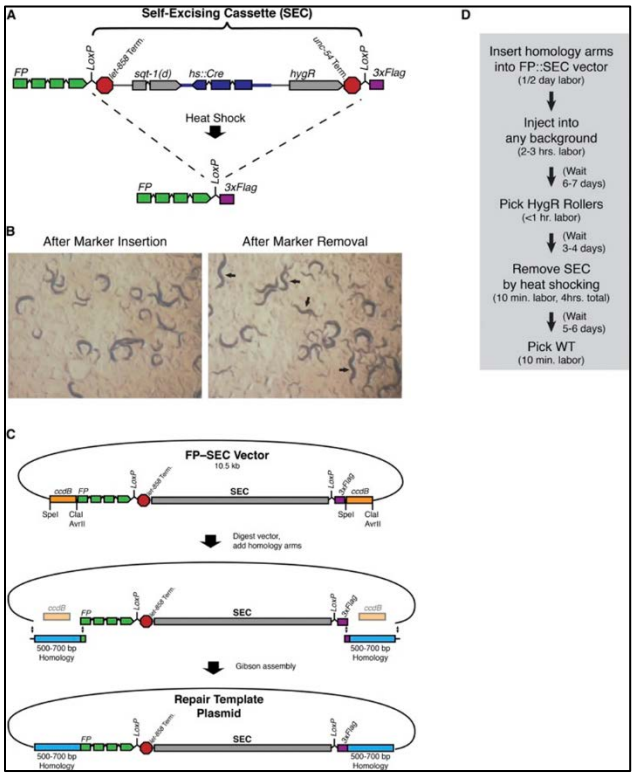


**Figure 5. Schematic for GFP insertion and worm selection (Dickinson and Goldstein, 2016).** Visual representation of the selection process and injection timeline used to generate worm strains containing the desired tagged gene within their chromosomes.

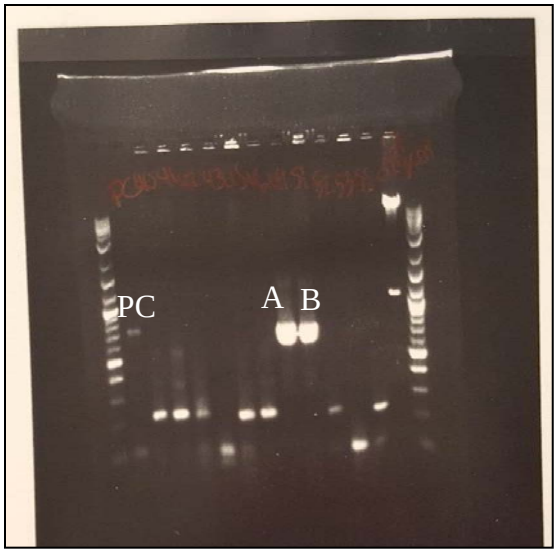


**Figure 6. Insertion genotyping for transformed cells containing *hmg-3* and Myc tag.** *PC.* Positive control for cells with appropriate insertion. *A.* and *B.* Samples from colonies where cells successfully included the gene and tag. Out of 11 colonies screened for the correct insertion, two (*A* and *B*) had the expected weight matching the positive control and indicating success of the process.

**Aim 2: Microscopy**

Using fluorescent microscopy conditions, the target proteins were visible *in vivo*, revealing the specific areas of the organism where the subunits were present. The location of the product in the cells of *C. elegans* differed between the HMG-3 and HMG-4 proteins (see Figures 7 and 8). HMG-3 had somatic presence during the embryonic stage, but in later larval stages the product was confined to the germline. In contrast, HMG-4 and SPT-16 were present in the somatic and germ cells throughout the entire developmental process of the organisms (see Figures 8 and 9).

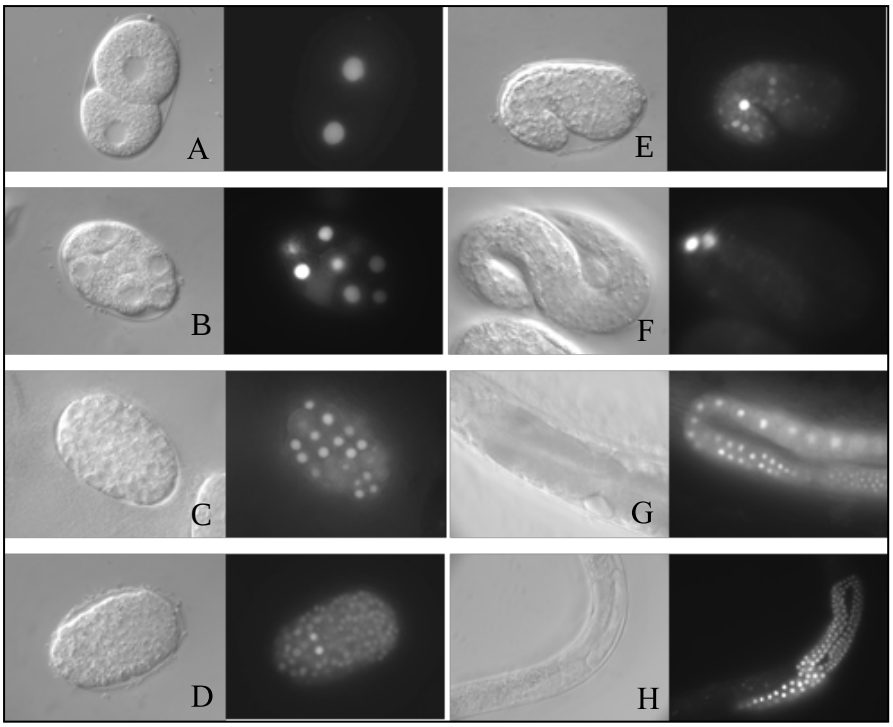


**Figure 7. HMG-3 GFP: presence of *hmg-3* protein product under fluorescent conditions in *C. elegans* throughout developmental time.** The protein is localized to the nuclei at all stages. *A.* Embryo in the 2-cell stage. *B.* Embryo undergoing early gastrulation. *C.* Embryo in late gastrulation. *D.* "Bean" stage of embryo; enclosure prior to elongation. *E.* Beginning of elongation. *F.* Complete elongation prior to hatching. The patches of brightest fluorescence represent the germ cells of the embryo at this stage. Very little somatic expression is present. *G.* Gonads of an adult hermaphrodite. *H.* Gonads of an adult male.

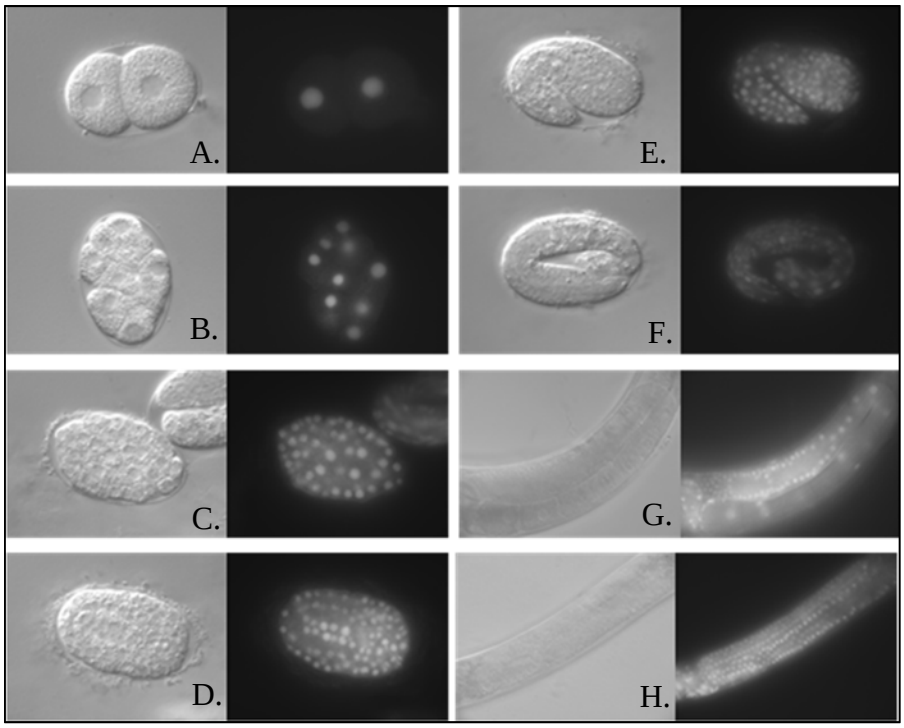


**Figure 8. HMG-4 GFP: presence of *hmg-4* protein product under fluorescent conditions in *C. elegans* throughout developmental time.** The protein is localized to the nuclei at all stages. *A.* Embryo in the 2-cell stage. *B.* Embryo undergoing early gastrulation. *C.* Embryo in late gastrulation. *D.* "Bean" stage of embryo; enclosure prior to elongation. *E.* Beginning of elongation. *F.* Complete elongation prior to hatching. *G.* Gonads of an adult hermaphrodite. *H.* Gonads of an adult male.

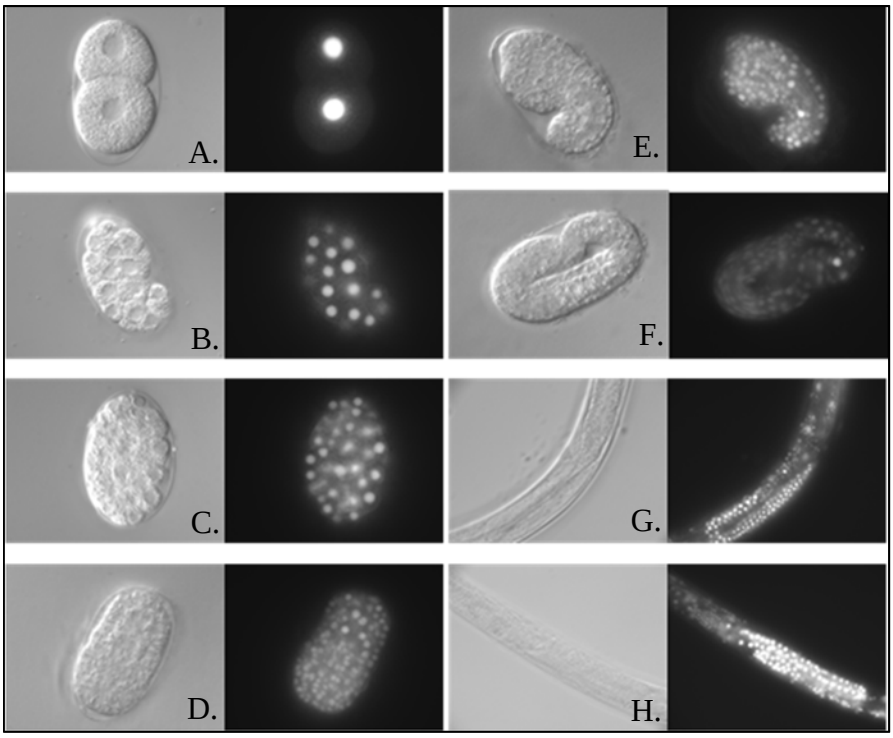


**Figure 9. SPT-16 GFP: presence of *spt-16* protein product under fluorescent conditions in *C. elegans* throughout developmental time.** The protein is localized to the nuclei at all stages. *A.* Embryo in the 2-cell stage. *B.* Embryo undergoing early gastrulation. *C.* Embryo in late gastrulation. *D.* "Bean" stage of embryo; enclosure prior to elongation. *E.* Beginning of elongation. *F.* Complete elongation prior to hatching. *G.* Gonads of an adult hermaphrodite. *H.* Gonads of an adult male.

**Aim 2: Western Blot**

Western Blot procedures against the FLAG tag for the proteins of interest resulted in product bands in the correct locations when compared to the protein ladder, confirming the presence of the intended FLAG tags in HMG-3, HMG-4, and SPT-16 (see Figure 10).

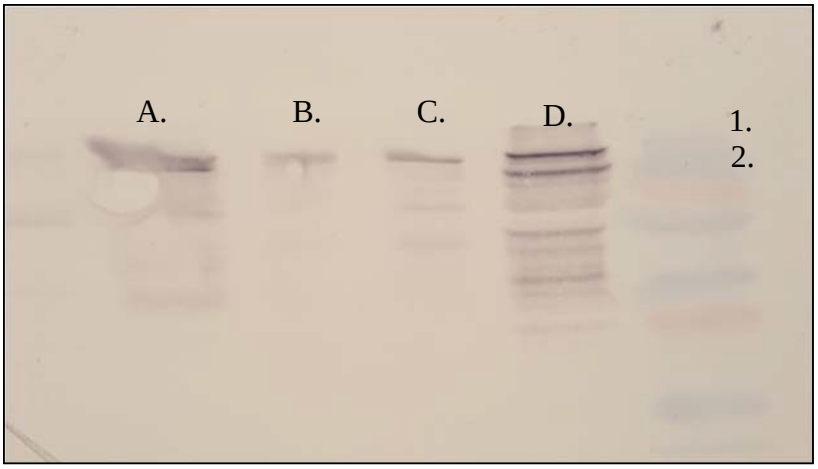


**Figure 10. Western Blot for Protein Tagging Confirmation.** *A.* HMG-3 with expected product band at 111 kDA. *B.* HMG-4-A, one of two HMG-4 strains that were successful with expected product band at 112 kDa. *C.* HMG-4-B, one of two HMG-4 strains that were successful with expected product band at 112 kDa. *D.* SPT-16 with expected product band at 150 kDa. *1.* Fermentas protein ladder mark for 130 kDa. *2.* Fermentas protein ladder mark for 100 kDa.

**DISCUSSION:**

HMG-4 and SPT-16 were found in both somatic and germline cells throughout the developmental process of the organism, which was to be expected as the FACT complex has important roles in DNA transcriptional processes. However, the HMG-3 proteins were found only in the germline cells and not in the somatic cells. In *C. elegans,* cells in the germline develop using proteins translated from the mRNA of the mother, and thus are only affected by the maternal genotype, without any contribution from the zygote (Alberts et al., 2002). Although *hmg-3* and *hmg-4* both encode the SSRP1 protein of the FACT complex, the expression of the product proteins of the genes is not what would be expected.

According to the duplication-degeneration-complementation (DDC) model of gene duplication, duplicate gene preservation, as is seen in this case with *hmg-3* and *hmg-4,* may arise when functions of an original ancestral gene are split between two genes via degenerative mutations over evolutionary time, and both genes are required to fulfill the function of the original gene

(Force et al., 1999). This process is called subfunctionalization. Interestingly, *hmg-3* and *hmg-4* can compensate for each other during embryonic development, but not larval development. Only one SSRP1 ortholog is required for embryonic survival, but in order to produce a fertile adult worm, both genes are necessary (Suggs et al., 2018). HMG-4 appears to have additional functions involving fertility (Suggs et al., 2018).

*Future Directions*

This exploratory study leaves many avenues for future research. One potential direction would be discovering why *hmg-3* and *hmg-4* have many overlapping functions but cannot completely compensate for each other in terms of complete development in both embryonic and larval stages. Future research may also include interactions between subunits and determining the number of each subunit protein in the FACT complex The FACT complex is highly conserved across many species, including humans, and with its potential HIV and cancer treatment applications, a more comprehensive understanding of the complex within eukaryotes is valuable.

## References


Alberts, B. et al. 2002. Molecular Biology of the Cell. 4th edition. New York: Garland Science. Caenorhabditis Elegans: Development from the Perspective of the Individual Cell. www.ncbi.nlm.nih.gov/books/NBK26861/ Accessed Mar. 9, 2018.

Dickinson, D.J., and Goldstein, B. 2016. CRISPR-based methods for *Caenorhabditis elegans* genome engineering. Genetics. 202:885-901. doi: 10.1534/genetics.115.1821162 Accessed Jan. 5, 2018.

Force, A. et al. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 151:1531-1545. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1460548/pdf/10101175.pdf Accessed Mar. 16, 2018.

Huang, H. et al. 2016. FACT proteins, SUPT16 and SSRP1, are transcriptional suppressors of HIV-1 and HTLV-1 that facilitate viral latency. The Journal of Biological Chemistry. doi: 10.1074/jbc.M115.652339 Accessed Mar. 20, 2018.

Koelle, M. n.d. Western Blot. https://medicine.yale.edu/lab/koelle/protocols/Western%20Blots_180546_21947_v1.pdf Accessed Mar. 18, 2018.

Kruger, A.V. et al. 2015. Comprehensive single cell-resolution analysis of the role of chromatin regulators in early *C. elegans* embryogenesis. Developmental Biology. 398:153-162. doi: 10.1016/j.ydbio.2014.10.014 Accessed Jan. 5, 2018.

NCBI. n.d. Align sequences nucleotide BLAST. https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=align2seq Accessed Mar. 23, 2018.

New England BioLabs. n.d. Monarch plasmid DNA miniprep kit protocol. https://www.neb.com/protocols/2015/11/20/monarch-plasmid-dna-miniprep-kit-protocol-t1010 Accessed Mar. 27, 2018.

Suggs, B.Z. et al. 2018. FACT complex genes have critical maternal and zygotic functions in *C. elegans.* Submitted to Developmental Biology for review: under revision.

Winkler, D.D., and Luger, K. The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization. The Journal of Biological Chemistry. 2011. doi: 10.1074/jbc.R110.180778 Accessed Jan. 5, 2018.